The genetics of smoking initiation and quantity smoked in Dutch adolescent and young adult twins

Koopmans, J.R.; Slutske, W.S.; Heath, A.C.; Neale, M.C.; Boomsma, D.I.

published in
Behavior Genetics
1999

DOI (link to publisher)
10.1023/A:1021618719735

document version
Publisher's PDF, also known as Version of record

Behavior Genetics, Vol. 29, No. 6, 1999

The Genetics of Smoking Initiation and Quantity Smoked in Dutch Adolescent and Young Adult Twins

Judith R. Koopmans,[1,5] Wendy S. Slutske,[2] Andrew C. Heath,[3] Michael C. Neale,[4] and Dorret I. Boomsma[1]

Received 2 May 1999—Final 17 Nov. 1999

Not much is known about the genetic and environmental determinants of various aspects of substance use in adolescents. This study examined whether the inheritance of initiation of tobacco use in adolescents is independent of the inheritance of the number of cigarettes smoked. Alternative multifactorial threshold models were applied to data on tobacco use in 1676 Dutch adolescent twin pairs. The three models that were considered are (i) the single liability dimension model, (ii) the independent liability dimension model, and (iii) the combined model (CM). The results showed that there is not one underlying continuum of liability to smoking. The CM was the best-fitting model. This model postulates that there are separate initiation and quantity dimensions but allows for the possibility that there are some individuals who are so low on the liability to level of consumption that they are not using tobacco. There were no differences between males and females in the magnitude of the genetic and environmental influences on individual differences in smoking initiation and quantity smoked. Smoking initiation was influenced by genetic factors (39%) and shared environmental influences (54%). Once smoking is initiated genetic factors determine to a large extent (86%) the quantity that is smoked.

KEY WORDS: Smoking initiation; quantity smoked; genetic influences; Dutch twins; adolescents; young adults.

INTRODUCTION

Evidence from large-scale population-based twin studies suggests that genetic factors contribute to individual differences in drinking behavior and smoking (e.g., Heath, 1995; Heath and Madden, 1995). An important issue, that most twin studies have not addressed, is whether the same or different genetic and environmental factors influence various aspects of substance use. For example, are smoking initiation and number of cigarettes smoked part of the same continuum of liability to smoking or are there independent genetic and environmental factors that determine initiation and the quantity smoked? It is important to understand the determinants of different aspects of substance use because an incorrect definition of the phenotype can lead to biased estimates of the genetic and environmental factors (Heath *et al.*, 1991a). If the same genetic and environmental factors determine whether or not a person is a smoker and how much is smoked, then exclusion of nonsmokers can lead to truncation of the distribution. In twin data this will lead to biased estimates of the heritability (Heath *et al.*, 1991a; Neale *et al.*, 1989). If the determinants of smoking initiation are independent of the determinants of number of cigarettes smoked, then inclusion of nonsmokers in the analyses of quantity measures may confound two different modes of inheritance (Heath *et al.*, 1991a).

Heath *et al.* (1991b) proposed three alternative multifactorial threshold models to test different as-

[1] Department of Biological Psychology, Vrije Universiteit, De Boelelaan 1111, 1081 HV Amsterdam, The Netherlands.
[2] Department of Psychology, University of Missouri, Columbia, Missouri; 65211.
[3] Department of Psychiatry, Washington University School of Medicine, St. Louis, Missouri 63110.
[4] Department of Human Genetics, Virginia Commonwealth University, Richmond, Virginia 23284.
[5] To whom correspondence should be addressed. Fax: +31.20.444.8832. e-mail:jr.koopmans@psy.vu.nl.

0001-8244/99/1100-0383$16.00/0

sumptions about the determinants of initiation and quantity of alcohol use. Similar models were used to test whether the inheritance of smoking initiation was independent of the inheritance of smoking persistence (Heath and Martin, 1993) (formally, despite use of the same labels, the "combined" models used in the two papers make somewhat different assumptions). Briefly, these models are (i) single liability dimension (SLD), which assumes that the same genetic and environmental factors influence initiation and quantity of substance use, but to a different degree; (ii) independent liability dimension (ILD), which assumes that the genetic and environmental determinants of initiation of substance use are separate from the determinants of quantity consumed; and (iii) combined (CM), which postulates that there are separate initiation and quantity dimensions but allows for the possibility that there are some individuals who are so low on the liability to level of consumption that they are not using substances. Until now the application of these models required purpose-written software. In this paper we used Mx (Neale, 1995) to fit the multifactorial threshold models to contingency tables by method of maximum likelihood. We apply the three multifactorial threshold models to data on initiation and quantity of tobacco use in a population-based sample of Dutch adolescent twins.

Heath *et al.* (1991b) applied the models described above to data on smoking initiation and smoking persistence in two cohorts of Australian twins (Heath and Martin, 1993). For the older cohort (aged 31 years and older) the independent liability dimension model gave the best fit to the data. The genetic effect on smoking persistence, explaining 53% of the total variance, was independent of the genetic and environmental effects on smoking initiation. For the young cohort (aged 18–30 years), Heath and Martin (1993) showed that the combined model gave the best description of the data, indicating that there were some genetic and environmental factors which influenced both smoking initiation and smoking persistence and other factors which influenced only persistence. In our twin-family study of health-related behavior we have found previously that individual differences in initiation of adolescent tobacco use could be attributed to shared environmental influences and small to moderate genetic influences (Boomsma *et al.*, 1994). There are three other twin studies that have assessed adolescent smoking (Han *et al.*, 1999; Hopper *et al.*, 1992; Maes *et al.*, 1999). So far, none of these studies reported the inheritance of the level of tobacco consumption. In this paper the question is addressed whether the inheritance of initiation of tobacco use in adolescents is independent of the inheritance of quantity smoked. After identification of the correct liability model, the relative contribution of genetic and environmental factors to initiation and quantity of tobacco use will be estimated.

METHODS

Sample

This study is part of an ongoing twin family study on health-related behavior in a population based sample of Dutch adolescent and young adult twins (Boomsma *et al.*, 1994; Koopmans and Boomsma, 1996; Koopmans *et al.*, 1995). The data were collected in 1991 from the first questionnaire on health and lifestyle that was mailed to 2375 adolescent twins and their parents. Completed questionnaires were returned by 1700 families. The age of the twins at the time of completing the questionnaire was between 12 and 24 years; the mean age was 17.7 years (SD = 2.26 years). Thirty-two percent of the twins were 12–15 years old, 22% were 16–17 years old, and 46% of the twins were 18 years of age and older. Less than 4% of the sample was younger than 14 years and 7% was older than 21. The sample of twins came from all regions of The Netherlands, including both rural and urban areas. In addition, the sample was representative of the general population of the Netherlands with regard to the educational level of the parents: 13.9 and 15.3% of the fathers and mothers, respectively, had a basic education at the elementary school (compared to 16.9 and 21.5% of similarly aged men and women, respectively, in the general population), 61.8 and 72.5% of the fathers and mothers, respectively, had a high school education (compared to 60.6 and 64.4% of similarly aged men and women, respectively, in the general population), and 24.3 and 12.1% of the fathers and mothers, respectively, had attained a college or university-level degree [compared to 22.5 and 13.9% of similarly-aged men and women, respectively, in the general population [(Koopmans, 1997)].

Zygosity of the twins was determined by questionnaire items about physical similarity and frequency of confusion of the twins by family and strangers (Goldsmith, 1991; Magnus *et al.*, 1983). The sample was divided into five groups by sex and zygosity of the twins; 275 pairs of monozygotic males (MZM), 360 monozygotic female twins (MZF), 259 dizygotic male twins (DZM), 322 dizygotic female twins (DZF), and 485 dizygotic opposite-sex twins (DOS). There

were 1676 twin pairs that provided complete data on smoking.

Measures

Smoking-initiation was assessed with the questions "Have you ever smoked?" and "Are you a smoker?" Current smokers (17%) and former smokers (7.7%) were asked how many cigarettes, cigars, or pipes on average they (had) smoked per day. Less than 1% of the sample reported that they smoked cigars or pipe. The twins were classified as nonsmokers if they answered no to the first question. Among 12- to 14-year-old males and females, 8.9 and 8.8% were smokers, respectively; among 15–16 year olds, 16.8 and 20.5% of males and females were smokers, respectively; and among 17–25 year olds, 38.2 and 32.2% of males and females were smokers, respectively (Koopmans *et al.*, 1997).

Current and former smokers were classified according to the daily amount of cigarettes smoked. For each zygosity group two-way contingency tables were computed. To avoid empty cells, the quantity data were collapsed into three categories. Smokers were divided into heavy smokers (>10 cigarettes per day), moderate smokers (6–10 cigarettes per day), and light smokers (1–5 cigarettes per day). Overall, 27.8% of individuals who ever smoked were classified as heavy smokers, 27.0% as moderate smokers, and 45.2% as light smokers.

Smoking in The Netherlands Versus the United States

The prevalences of lifetime tobacco use initiation among individuals 12 years of age and older in large general population surveys conducted in 1997 in The Netherlands (Abraham *et al.*, 1999) and the United States (Office of Applied Studies, 1999) were 67.9 and 70.5, respectively. The annual average consumptions of cigarettes among adults in The Netherlands and the United States are also quite similar (World Health Organization, 1997).

The Netherlands and the United States differ with respect to official policies concerning the selling of cigarettes to minors, in that there is no law restricting the sale of cigarettes to minors in The Netherlands, whereas in the United States, one must be at least 18 years of age to purchase cigarettes legally. However, most youth in the United States feel that it would be fairly easy to obtain cigarettes if they wanted them (Johnston *et al.*, 1998), so it is unclear whether such laws are effective at deterring adolescent smoking. Neither country allows advertising of tobacco products on radio and television, and in The Netherlands advertising of tobacco products on billboards is also prohibited. Perhaps the greatest deterrent of cigarette smoking among youth in The Netherlands is the price—taxes account for about 72% of the price of cigarettes in The Netherlands, compared to about 30% of the price in the United States. Thus, the price of a pack of cigarettes in The Netherlands is roughly 40% higher than in the United States (World Health Organization, 1997).

Despite these differences, it appears that the epidemiologies of smoking among youth in The Netherlands and the United States are roughly equivalent. In The Netherlands, 35.3% of 12–15 year olds; 58.0% of 16–19 years olds, and 59.4% of 20–24 years olds have ever smoked (Abraham *et al.*, 1999). In the United States, 38.7% of 12–17 year olds and 67.7% of 18–25 year olds have ever smoked (Office of Applied Studies, 1999). The mean age of first tobacco use is 17 years in The Netherlands, and by age 16, 50% of all lifetime tobacco users have initiated use (Abraham *et al.*, 1999).

Liability Models

Figure 1 shows three models for the relationship between the genetic and the environmental determinants of initiation and quantity of substance use. Heath *et al.* (1991b) postulated these three models to describe genetic and environmental influences on abstinence, frequency, and quantity of alcohol consumption in adult Australian twins. In this paper these models are applied to tobacco use. The single liability dimension (SLD) model (Fig. 1a) assumes that the liability to smoking is unidimensional and is normally distributed and determines both initiation and quantity of tobacco use. Under this model the same genetic and environmental risk factors predispose to smoking initiation and to quantity smoked. Heavy smokers are influenced by more extreme genetic or environmental factors than light smokers or nonsmokers. The underlying normal liability distribution is divided by thresholds into discrete categories, which, in the case of the SLD model, correspond to the observed categories. Individuals falling between threshold t_0 and t_1 will be heavy smokers, those falling between t_1 and t_2 will be moderate smokers, etc. The probability that an individual falls into one of the four categories is given by y_1, y_2, y_3, and y_4 in Fig. 1a and can be calculated by integrating a standardized normal distribution between the corresponding threshold values. The model predicts that cotwins

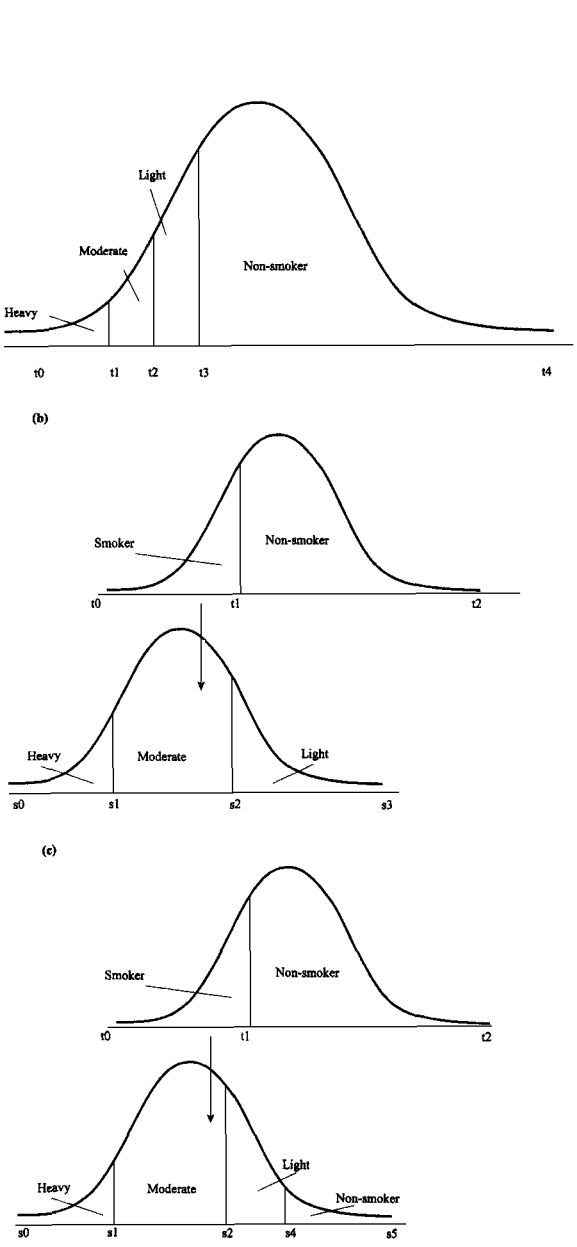

Fig. 1. Normal liability distribution for smoking under the single liability dimension (SLD) model (a) and normal liability distributions for initiation and quantity of tobacco use under the independent liability dimension (ILD) model (b) and the combined model (CM) (c).

of heavy smoking twins are more likely to be heavy smokers than are cotwins of light smokers.

The ILD model (Fig. 1b) postulates two independent liability dimensions for initiation and quantity that are each determined by completely separate genetic and environmental factors. The initiation dimension determines whether or not a person will become a smoker. Individuals falling below the threshold t_1 are predicted to be smokers. The quantity dimension determines whether an individual becomes a heavy, moderate, or light smoker, given that he/she is a smoker, with conditional probabilities x_1, x_2, and x_3 (see Fig. 1b). The probabilities that an individual will be a heavy smoker, moderate smoker, light smoker, or nonsmoker are y_1x_1, y_1x_2, y_1x_3, and y_2, respectively, where y_1 and y_2 are the unconditional probabilities for initiating smoking and remaining a nonsmoker, respectively. The ILD model predicts that the cotwin of a nonsmoking twin is also more likely a nonsmoker, but if the cotwin of a nonsmoking twin is a smoker, he/she will not, on average, differ in the amount of cigarettes smoked from smoking cotwins of smoking twins.

The CM (Fig. 1c) includes features of both the SLD and the ILD models. The SLD and ILD models are nested under the more general CM. Like the ILD model, the CM allows for independent initiation and quantity dimensions, with different genetic and environmental factors that determines whether or not a person is a smoker and the total amount of cigarettes smoked. However, like the SLD model, those on the quantity dimension can become nonsmokers due to low exposure to risk factors which influence the quantity smoked. Thus, under the CM there are two routes to nonsmoking. Under the CM the cotwin of a smoking twin is more likely to become a nonsmoker than under the ILD model.

Model Fitting

Tobacco use (heavy, moderate, light, or abstaining) in the first twin was cross-classified with tobacco use in the second twin, resulting in 4×4 contingency tables for each zygosity group (see also Table I). Models were fitted to the five contingency tables by the method of maximum likelihood with Mx (Neale, 1995).

The analyses were based on the assumption that the observed discrete distribution (i.e., heavy, moderate, light, or nonsmoker) has an underlying continuous distribution that has been termed the liability (Falconer, 1989). Thresholds divide this normal liability distribution into discrete categories. The joint distributions of twin pairs for the liability dimensions are assumed to be bivariate normal, with correlation r_i being the correlation in liability between twins for the ith zygosity group. For each of the liability dimensions the polychoric twin pair correlations and the thresholds were estimated by maximum likelihood. The thresholds were allowed to be different for males

Table I. Predicted Probabilities for a Twin Pair under the Single Liability Dimension (SLD), the Independent Liability Dimension (ILD) and the Combined Model (CM)[a]

		Twin 2			
Twin 1	Model	Heavy	Moderate	Light	Nonsmoker
Heavy	SLD	y_{11}	y_{12}	y_{13}	y_{14}
	ILD	$y_{11}x_{11}$	$y_{11}x_{12}$	$y_{11}x_{13}$	$y_{12}x_{1.}$
	CM	$y_{11}x_{11}$	$y_{11}x_{12}$	$y_{11}x_{13}$	$y_{11}x_{14} + y_{12}x_{1.}$
Moderate	SLD	y_{21}	y_{22}	y_{23}	y_{24}
	ILD	$y_{11}x_{21}$	$y_{11}x_{22}$	$y_{11}x_{23}$	$y_{12}x_{2.}$
	CM	$y_{11}x_{21}$	$y_{11}x_{22}$	$y_{11}x_{23}$	$y_{11}x_{24} + y_{12}x_{2.}$
Light	SLD	y_{31}	y_{32}	y_{33}	y_{34}
	ILD	$y_{11}x_{31}$	$y_{11}x_{32}$	$y_{11}x_{33}$	$y_{12}x_{3.}$
	CM	$y_{11}x_{31}$	$y_{11}x_{32}$	$y_{11}x_{33}$	$y_{11}x_{34} + y_{12}x_{3.}$
Nonsmoker	SLD	y_{41}	y_{42}	y_{43}	y_{44}
	ILD	$y_{21}x_{.1}$	$y_{21}x_{.2}$	$y_{21}x_{.3}$	y_{22}
	CM	$y_{11}x_{41}$ + $y_{21}x_{.1}$	$y_{11}x_{42}$ + $y_{21}x_{.2}$	$y_{11}x_{43}$ + $y_{21}x_{.3}$	$y_{11}x_{44} + y_{21}x_{.4}$ + $y_{12}x_{4.} + y_{22}$

[a] Under the SLD model, y_{jk} = the probability that a twin pair falls in the j,k-th category of smoking. Under the ILD and combined model, y_{jk} = the probability that a twin pair falls in the j,k-th category of the initiation dimension; x_{jk} = the probability that a twin pair falls in the j,k-th category of the quantity dimension; $x_{j.}$ = the probability that the first twin falls in the j-th category of the quantity dimension; $x_{.k}$ = the probability that the second twin falls in the k-th category of the quantity dimension.

and females. Under the SLD model one twin correlation for each zygosity group and three thresholds for males (t_1, t_2, t_3) and females (t'_1, t'_2, t'_3) were estimated, with $t_0 = t'_0 = -\infty$ and $t_4 = t'_4 = \infty$ (see Fig. 1a), giving in total 11 parameters to be estimated. Under the ILD model and CM separate twin correlations for the initiation and quantity dimensions were estimated for each zygosity group. For the initiation dimension two thresholds (t_1 and t'_1) were estimated, with $t_0 = t'_0 = -\infty$ and $t_2 = t'_2 = \infty$. Under the ILD model there was no abstinence category for the quantity dimension, leaving four thresholds to be estimated (s_1, s_2, s'_1, s'_2), with $s_0 = s'_0 = -\infty$ and $s_3 = s'_3 = \infty$. Under the CM the same number of thresholds was estimated for the quantity dimension as for the SLD model (three thresholds for males and three for females). There were 16 parameters to be estimated under the ILD model, and 18 parameters under the CM.

The probability that a twin pair from the ith zygosity group falls into the j,k-th cell of the ith contingency table is calculated by

$$y(i,j,k) = \phi(t_j, t_k) - \phi(t_{j-1}, t_k) - \phi(t_j, t_{k-1}) + \phi(t_{j-1}, t_{k-1}) \quad (1)$$

where $\Phi(t_j, t_k)$ represents the integrated bivariate normal density from $-\infty$ to t_j and from $-\infty$ to t_k with correlation r_i between twins. Equation (1) gives the unconditional probability $y(i,j,k)$ for smoking (SLD model) or for smoking initiation (ILD model and CM). For the quantity dimension of the ILD model and CM the conditional probability $x(i,j,k)$ can be obtained by

$$x(i,j,k) = \phi(s_j, s_k) - \phi(s_{j-1}, s_k) - \phi(s_j, s_{k-1}) + \phi(s_{j-1}, s_{k-1}) \quad (2)$$

where $\Phi(s_j, s_k)$ represents the integrated bivariate normal density from $-\infty$ to s_j and from $-\infty$ to s_k with correlation r'_i, where r'_i is the twin correlation for liability to quantity. The predicted probabilities for a twin pair under the three models are given in Table I. Under the SLD model, y_{11} denotes the probability that both twins are heavy smokers, y_{12} denotes the probability that the first twin is a heavy smoker and the second twin is a moderate smoker, and so on. Under the ILD model and CM, y_{11}, y_{22}, y_{12}, and y_{21} denote the probabilities that twins both fall into the smoking category, both fall into the abstinent category, or are discordant for smoking status at the initiation dimension. The conditional probability that both twins are heavy smokers, the first twin is a heavy smoker and the second twin is a moderate smoker, etc., is represented by x_{11}, x_{12}, etc.; x_j, denotes the probability that the first twin falls into the jth category of the quantity dimension, and x_k denotes the probability that the second twin falls into the kth category of the quantity dimension. Under the CM there are two routes to nonsmoking. For example, $y_{11}x_{14} + y_{12}x_1$ gives the probability that both twins are smokers on the initiation dimension (y_{11}) and the first twin is a heavy smoker while the second twin is a nonsmoker on the quantity dimension (x_{14}) plus the probability that the first twin is a smoker and the second twin is a nonsmoker on the initiation dimension (y_{12}) and the first twin is a heavy smoker (x_1).

Let $p(i,j,k)$ denote the probability, under a given model, that a twin pair from the ith zygosity group will fall in the j,k-th cell of the ith contingency table. Under the SLD model, $p(i,j,k) = y(i,j,k)$ in Eq. (1) for all i,j,k. Under the other two models, $p(i,j,k)$ is the predicted probability as given in Table I. The log-likelihood of a set of observations, under a given model, is given by

$$LL = \ln(c) + \sum\sum\sum f(i,j,k) \ln(p(i,j,k)) \quad (3)$$

where c is a constant, and $f(i,j,k)$ is the observed frequency of twin pairs from the ith twin group in the j,k-th cell of the observed contingency table. Maximum-likelihood estimates of the model parameters are obtained by maximizing this function with respect to the parameter values. The goodness of fit of nested models was assessed with likelihood-ratio chi-square tests.

Genetic Models

The three models were fitted to the data, estimating separate polychoric correlations for each zygosity group. For the model that gave the best description of the data, the twin correlations in liability were expressed as a function of genetic and environmental parameters based on the classical twin design (Neale and Cardon, 1992). For both the initiation and the quantity dimension, different genetic models were fitted. Under the full model (ACE), both additive genetic and shared environmental factors contribute to resemblances between twins. Sex differences were tested by allowing the magnitude of the genetic and environmental effects to be different for males and females and by allowing the correlation between the shared environmental factors or the genetic factors in opposite-sex twins to be less than unity. If the phenotypic correlation in opposite-sex twins is lower than the same-sex dizygotic twin correlations, this might be due to shared environmental effects that influence one sex but not the other or genetic effects that are expressed in one sex but not in the other. If both additive genetic and shared environmental factors contribute substantially to individual differences in both males and females, it is not possible to distinguish between these two effects with twin data (Eaves, 1977). Under the additive genetic (AE) model individual differences are explained by additive genetic influences and by environmental effects that are unique for an individual. Under the shared environmental (CE) model individual differences are explained by environmental influences that are shared between family members and by individual-specific environmental factors. For all models, different thresholds were estimated for males and females, allowing for differences in the prevalence of substance use between males and females.

RESULTS

Table II shows the cross-classification of tobacco use in the first twin with tobacco use in the second twin. For opposite-sex twins the data were reordered so that tobacco use in male twins was cross-classified with tobacco use in the female cotwins. Table II also shows the proportions of heavy smokers, moderate smokers, light smokers, and nonsmokers for first- and second-born twins in each zygosity group. There are no sex

Table II. Twin Concordance for Number of Cigarettes Smoked per Day with Proportions for First and Second Twins of Heavy, Moderate, Light or Nonsmokers

		Females					Males				
Twin 1	Twin 2	>10	6–10	1–5	non		>10	6–10	1–5	non	
MZ		$n = 355$					$n = 272$				
	>10	11	3	2	3	5.4%	11	1	2	5	7.0%
	6–10	4	7	3	3	4.8%	4	8	2	4	6.6%
	1–5	1	3	29	12	12.7%	1	3	12	11	9.9%
	nonsmoker	2	4	12	256	77.2%	3	4	7	194	76.5%
		5.1%	4.8%	13.0%	77.2%		7.0%	5.9%	8.5%	78.7%	
DZ		$n = 315$					$n = 252$				
	>10	9	2	6	5	7.0%	8	8	1	3	7.9%
	6–10	1	1	1	8	3.5%	4	4	5	8	8.3%
	1–5	4	8	16	10	12.1%	2	1	5	21	11.5%
	nonsmoker	6	9	19	210	77.5%	7	6	8	161	72.2%
		6.3%	6.3%	13.3%	74.0%		8.3%	7.5%	7.5%	76.6%	
DOS	female	$n = 482$									
male	>10	11	9	7	14	8.5%					
	6–10	5	5	8	28	9.5%					
	1–5	4	5	17	24	10.4%					
	nonsmoker	11	20	23	291	71.6%					
		6.4%	8.1%	11.4%	74.1%						

Table III. Estimated Polychoric Twin Pair Correlations (r) with 95% Confidence Intervals (CI) for the Initiation and Quantity Dimensions under the Full Combined Model and Estimated Thresholds for Males and Females[a]

	Correlations				Thresholds			
	initiation		quantity		initiation		quantity	
	r	95%CI	r	95%CI	tl	sl	s2	s3
MZM	0.91	0.68–0.98	0.84	0.52–0.96	-0.44	-0.71	-0.05	0.78
DZM	0.76	0.39–0.99	0.65	0.17–0.89				
MZF	0.94	0.84–0.99	0.88	0.69–0.96	-0.55	-0.82	-0.25	0.94
DZF	0.82	0.59–1.00	0.44	-0.09–0.79				
DOS	0.64	0.40–0.71	0.47	-0.02–0.78				

[a] MZM = monozygotic male twins; DZM = dizygotic males; MZF = monozygotic females; DZF = dizygotic females; DOS = dizygotic opposite-sex twins.

Table IV. Goodness-of-Fit of the Single Liability Dimension (SLD), the Independent Liability Dimension (ILD) and the Combined Model (CM) to the Data on Tobacco Use

		Smoking	
Model	df	χ^2	p
SLD	64	116.87	<0.001
ILD	59	78.79	0.04
CM	57	67.68	0.16

differences in the proportion of smokers and nonsmokers or in the quantity smoked.

The three models of liability underlying the initiation and the quantity dimension were fitted to the data, estimating polychoric correlations for each zygosity group. Table III shows the estimated polychoric twin correlations for each zygosity group for the initiation and the quantity dimension under the full CM. The difference between the MZ and the DZ correlations for the initiation dimension suggested that both shared environmental and genetic factors are important, whereas the pattern of correlations for the quantity dimension suggested that genetic factors are more important.

Table IV gives the goodness of fit for each liability model. The SLD model for smoking was rejected. The combined model gave the best description of the data. Thus, a person can be a nonsmoker due to genetic and/or environmental factors that influence the initiation dimension or because that person is low on the quantity dimension that determines the number of cigarettes smoked per day. The predicted marginal probabilities for smoking under the full CM are represented in Fig. 2. The percentage of nonsmokers in the total sample that comes from the quantity dimension is 7% (= 0.33 × 0.22) in males (Fig. 2a) and 5% in females (Fig. 2b).

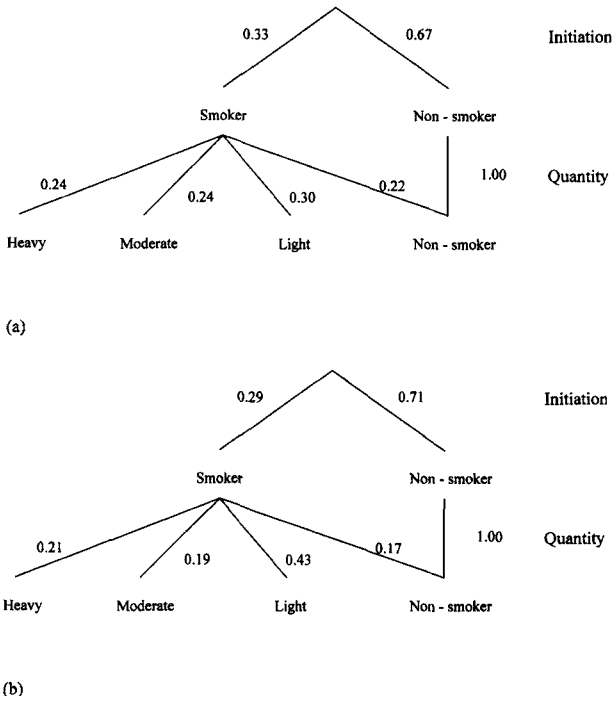

Fig. 2. Estimated probabilities under the full combined model for (a) smoking in males and (b) smoking in females.

Different genetic models were fitted both to the initiation dimension and to the quantity dimension under the combined model (Table V). The initiation dimension was best described by a model without sex differences and with both additive genetic and shared environmental effects (model 2). The correlation between the shared environmental factors in opposite-sex twins was estimated as 0.77 and could be constrained to unity without a significant reduction in the goodness of fit. For the quantity dimension the full model could be

Table V. Model Fitting Results for Smoking under the Combined Model (Best Fitting Model is Given in Boldface)[a]

Genetic model					
Initiation	Quantity	χ^2	df	p	AIC
1. Full	full	67.70	57	0.16	-46.30
2. ACE	full	70.85	60	0.16	-49.15
3. AE	full	86.13	61	0.02	-35.87
4. CE	full	75.60	61	0.10	-46.40
5. Full	ACE	69.24	60	0.19	-50.76
6. Full	AE	70.34	61	0.19	-51.66
7. Full	CE	84.38	61	0.03	-37.62
8. ACE	AE	**75.65**	**64**	**0.15**	**-52.35**

[a] Full = full model with sex-dependent effects and a correlation between the shared environmental factors in opposite-sex twins (= r_c) that is allowed to be less than 1; ACE = full model without sex differences; AE = additive genetic model; CE = shared environmental model. AIC = χ^2 - 2df, this is a measure of the parsimony of the model, a lower value of AIC indicates a more parsimonious model.

reduced to an additive genetic model without sex differences (model 6). The genetic correlation between males and females in opposite-sex twins was estimated as 0.5. Thus there was no evidence for different genetic factors in males and females influencing the quantity smoked. The best-fitting model was model 8, which fitted the full model without sex differences to the initiation dimension and the additive genetic model without sex differences to the quantity dimension. The parameter estimates for the best-fitting model are given in Table VI. The magnitudes of the genetic and environmental influences on smoking initiation and quantity smoked were the same for males and females. Individual differences in smoking initiation could be explained by shared environmental factors (54%) and moderate genetic influences (39%), whereas 86% of the total variation in the number of cigarettes smoked per day could be explained by genetic factors.

DISCUSSION

Why do adolescents start to smoke? Our results showed that for both males and females, shared environmental influences and moderate genetic influences contribute to variation in smoking initiation. Once adolescents start smoking, genetic factors explain to a large extent individual differences in the number of cigarettes smoked. The three multifactorial threshold models that were fitted showed that there is not one underlying continuum of liability to smoking. The CM was the best-

Table VI. Proportions of the Total Variance in Initiation and Quantity of Tobacco use that are Explained by Additive Genetic Factors (h^2), Shared Environmental Influences (c^2) and Unique Environmental Effects (e^2) under the Best Fitting Model (95% Confidence Intervals of the Parameter Estimates are Given Between Parentheses)

	h^2	c^2	e^2
Initiation	0.39 (.00–.68)	0.54 (.25–.95)	0.07 (.02–.16)
Quantity	0.86 (.70–.94)	—	0.14 (.06–.30)

fitting model. Under this model there are two routes to become a nonsmoker: a person does not smoke because of genetic and environmental factors that influence the liability to smoking initiation or because that person has a low exposure to the genetic and environmental risk factors which influence the quantity smoked.

Most studies on smoking have assessed adult smoking behavior. Heath and Madden (1995) reviewed the data from the major large-scale twin studies and concluded that there is an important genetic influence on all aspects of smoking behavior, including smoking initiation, amount smoked, and smoking persistence, but not much has been published about the genetics of adolescent smoking behavior. In a study of 1400 adolescent Australian twin pairs, aged 11 to 18 years, twin associations for smoking were represented by log odds ratios (Hopper *et al.*, 1992). Higher odds ratios in MZ twins compared to DZ twins suggested that genetic factors played a role in determining smoking in adolescents. Differences in odds ratios suggested stronger genetic influences in males compared to females, although this was not formally tested. In a sample of 571 17- to 18-year-old Minnesota-born twins, 44% of the variation in smoking initiation was explained by shared environmental factors and 36% of the variation was explained by genetic factors (Han *et al.*, 1999). Our results were similar to those of Han *et al.* (1999); in our sample of twins, aged 12–24 years (mean age, 17.7 years), 54% of the variation in smoking initiation was explained by shared environmental factors and 39% of the variation was explained by genetic factors. Seemingly contradictory results come from a study of 170 16-year-old twins in Virginia, in which only 18% of the variation in smoking initiation was explained by shared environmental factors and 65% of the variation was explained by genetic factors (Maes *et al.*, 1999). However, in the study by Maes *et al.* (1999), tobacco use was defined as "ever consuming more than 1 cigarette or other tobacco products per day," which is a higher

threshold than other studies have used to define smoking initiation. We suspect that minor differences in survey questions can result in substantial differences in estimates of the prevalence of and environmental and genetic contributions to smoking initiation. In the present study, the question "Have you ever smoked?" yielded a much lower prevalence of smoking initiation than has been obtained in other large epidemiological surveys of Dutch adolescents (Abraham *et al.*, 1999). Thus, we suspect that some of the adolescents in the present study who were infrequent smokers may have defined themselves as "nonsmokers." (Another explanation for the lower prevalence of smoking initiation in the present study is that smoking twins were less likely than nonsmoking twins to participate.)

Although there are differences in the estimates between different studies (Heath and Madden, 1995), smoking initiation is one of the few behavioral traits in adults for which an important contribution of shared environmental influences is found. We found in adolescent twins that shared environmental influences on smoking initiation were more important than genetic factors. What constitutes these shared environmental influences? Parent–offspring models showed that the shared environmental factor on smoking initiation is not influenced by parental smoking behavior (Boomsma *et al.*, 1994). The resemblances between parents and children for smoking could be explained by their genetic relatedness. The importance of shared environmental influences on smoking initiation parallels findings for alcohol use in adolescents (Heath, 1995; Heath and Martin, 1988; Koopmans and Boomsma, 1996), and there is evidence to suggest that the same shared environmental factors influence initiation of alcohol and tobacco use (Koopmans *et al.*, 1997). Peer influence is widely believed to be of major importance for the involvement of adolescents in smoking and drinking. This association may also reflect assortative friendship (Heath and Martin, 1988), which is the active selection of friends with similar behaviors. A longitudinal study of 1028 students controlled for the effects of friendship selection and found evidence for modest influences from the closest friend for initiation of cigarette and alcohol use (Urberg *et al.*, 1997). The friendship group use predicted transition into current cigarette use, whereas only the close friend use predicted transition into current alcohol use. Evidence from an adoption study by McGue *et al.* (1996) suggests that some of the shared environmental influences on adolescent alcohol use are due to sibling effects. Another aspect of the shared environment that has been shown to be inversely related to smoking and alcohol use in adolescents and adults is religious affiliation and religiosity (Heath and Martin, 1988; Kendler *et al.*, 1997). In our sample of adolescent twins, religious involvement of the twins was also found to be associated with smoking (Rietveld *et al.*, 1996). Adolescents who were actively involved were less likely to smoke than those who had a religious affiliation but were not actively involved.

Once smoking is initiated, genetic factors determine to a large extent (86%) whether an adolescent becomes a light, moderate, or heavy smoker. For quantity smoked no evidence was found for shared environmental influences. This is in contrast with our findings on alcohol use in the same sample of twins (Koopmans, 1997). Given that an adolescent is a drinker, the amount of alcohol consumption is influenced by genetic effects and shared environmental factors, which explain 32 and 44% of the variance in liability, respectively. Urberg *et al.* (1997) showed that both friendship group and best friends independently contributed to the prediction of adolescents' drinking to intoxication.

Are smoking initiation and quantity smoked part of the same liability to smoking? Heath and Martin (1993) proposed three multifactorial threshold models to account for the possibility that smoking initiation is a different dimension than smoking persistence. Heath and Martin (1993) found, in an cohort of Australian twins aged more than 30 years, that the substantial genetic influence on smoking persistence was independent of genetic effects on smoking initiation. In the younger cohort (aged 18–30 years) the combined model was the best-fitting model, suggesting that there are some genetic and/or environmental factors that influence both smoking initiation and smoking persistence. In this young cohort it was not possible to distinguish between genetic and nongenetic models to explain familial resemblances in smoking behavior. True *et al.* (1997) applied the same models to data on smoking persistence from the Vietnam Era Twin Registry. The male twins were between 30 and 48 years old at the time of the survey. Both genetic and shared environmental effects on smoking initiation were found, accounting for 50 and 30% of the variance, while only genetic factors influenced persistence in smoking (70%). As in the young cohort of Australian twins, it was found that there is a combined liability to smoking initiation and smoking persistence. In our study on adolescent twins the combined liability model also was the best-fitting model, suggesting that there are genetic and environmental factors that influence both smoking initiation and quantity smoked. In other words, there

are potential smokers on the initiation dimension who became nonsmokers due to their genetic predisposition on the quantity dimension. Although this was less than 10% of the sample of nonsmokers, this group might become larger when the twins grow older and all have passed the age of onset of risk. Kendler *et al.* (1999) developed a model that estimated the correlation between the liability to smoking initiation and the liability to nicotine dependence. In a population-based sample of adult female twins they found that the liabilities to smoking initiation and nicotine dependence were substantially correlated but not identical. Kendler *et al.* (1999) showed that there were some genetic factors that were specific for nicotine dependence, explaining 31% of the total genetic variance of nicotine dependence.

In contrast to an earlier report (Boomsma *et al.*, 1994), we did not find evidence for different shared environmental factors influencing smoking initiation in males and females. This difference in results from the same sample of adolescent twins is probably due to the different techniques that were used to analyze the data. In another report (Koopmans *et al.*, 1997), we also found substantial differences in the relative importance of shared environmental and genetic influences in the risk of smoking initiation for younger (12–16 years old) versus older (17–25 years old) twins, with genetic factors being relatively unimportant at the earlier ages and more important at the later ages. It was not possible to stratify the sample by age for the analyses in the present paper, but it is likely that there are important age differences in the role of shared environmental and genetic factors in smoking initiation, and so the estimate provided in the present paper should be considered the average contribution of shared environment and genes across age groups. It is not known whether there are important age differences in the role of shared environmental and genetic factors in quantity smoked (or smoking persistence). The results of the present study of adolescents and young adults, many of whom were not yet completely through the age of risk for smoking onset or maximal smoking involvement, are remarkably consistent with the results of the study by True *et al.* (1997), based on an adult sample that was through the age period of risk for smoking onset. This suggests that we would probably obtain similar results if we were to repeat these analyses when all of the subjects in the present study reach adulthood and have passed through the age period of risk for smoking onset.

One possible genetic mechanism that is involved in both smoking initiation and quantity smoked is impaired nicotine metabolism. Nicotine is metabolized to cotinine by the genetically variable enzyme CYP2A6. Recently it has been shown that individuals with impaired nicotine metabolism due to the lack of full functional CYP2A6 have a significantly reduced risk to become nicotine dependent (Pianezza *et al.*, 1998). In addition, among nicotine dependent smokers, those with impaired nicotine metabolism smoked significantly fewer cigarettes than smokers with two CYP2A6 active alleles. Sensation seeking and other heritable personality traits that are associated with adolescent smoking might also mediate the genetic influences on smoking initiation (Gilbert, 1995; Zuckerman, 1994). Futhermore, genetic differences in the sensitivity to nicotine, in the development of tolerance to nicotine, and in the rewarding effects of nicotine are most likely to be involved in the individual differences in the amount of tobacco use (Collins and Marks, 1991; Pomerleau, 1995). With genetic association and linkage studies, more will be learned about the biological mechanisms that are involved in smoking behavior.

In conclusion, our results show that there is not one underlying continuum of liability to smoking initiation and number of cigarettes consumed. For smoking initiation there is an important influence of shared environmental factors, while for quantity smoked only genetic factors are important. Some of the genetic factors that influence smoking initiation might also be involved with quantity smoked. Other studies have shown that both smoking persistence and nicotine dependence in adults are highly heritable. This study shows that even in adolescents, given that they are smokers, genetic factors determine to a large extent the number of cigarettes consumed. Future studies will have to show how quantity of tobacco use in adolescence is related to nicotine dependence in young adulthood and whether the same genetic and environmental factors are involved.

ACKNOWLEDGMENTS

This work was supported by the Netherlands Organization for Scientific Research (NWO Grant 904-61-072) and by the Free University (Grant USF96/22).

REFERENCES

Abraham, M. D., Cohen, P. D. A., van Til, R., and de Winter, M. A. L. (1999). *Licit and Illicit Drug Use in The Netherlands, 1997*, CEDRO Centrum Voor Drugsonderzoek, Universiteit van Amsterdam, Amsterdam.

Boomsma, D. I., Koopmans, J. R., van Doornen, L. J. P., and Orlebeke, J. F. (1994). Genetic and social influences on starting to smoke: A study of Dutch adolescent twins and their parents. *Addiction* **89**:219–226.

Collins, A. C., and Marks, M. J. (1991). Progress towards the development of animal related models of smoking-related behaviors. *J. Addict. Dis.* **10**:109–126.

Eaves, L. J. (1977). Inferring the causes of human variation. *J. Roy. Stat. Soc. Ser. A* **140**:324–355.

Falconer, D. S. (1989). *Introduction to Quantitative Genetics, 3rd ed.,* Longman Group, Essex.

Gilbert, D. G. (1995). *Smoking: Individual Differences, Psychopathology and Emotion,* Taylor and Francis, Washington, DC.

Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behav. Genet.* **21**:257–270.

Han, C., McGue, M. K., and Iacono, W. G. (1999). Lifetime tobacco, alcohol and other substance use in adolescent Minnesota twins: Univariate and multivariate behavioral genetic analyses. *Addiction* **94**:981–993.

Heath, A. C. (1995). Genetic influences on drinking behavior in humans. In Begleiter, H., and Kissin, B. (eds.), *Alcohol and Alcoholism, Vol 1. Genetic Factors and Alcoholism,* Oxford University Press, New York, pp. 82–121.

Heath, A. C., and Madden, P. A. F. (1995). Genetic influences on smoking behavior. In Turner, J. R., Cardon, L. R., and Hewitt, J. K. (eds.), *Behavior Genetic Approaches in Behavioral Medicine,* Plenum Press, New York, pp. 45–65.

Heath, A. C., and Martin, N. G. (1988). Teenage alcohol use in the Australian Twin Register: Genetic and social determinants of starting to drink. *Alcohol. Clin. Exp. Res.* **12**:735–741.

Heath, A. C., and Martin, N. G. (1993). Genetic models for the natural history of smoking: Evidence for a genetic influence on smoking persistence. *Addict. Behav.* **18**:19–34.

Heath, A. C., Meyer, J., Eaves, L. J., and Martin, N. G. (1991a). The inheritance of alcohol consumption patterns in a general population twin sample: I. Multidimensional scaling of quantity/ frequency data. *J. Stud. Alcohol* **52**:345–351.

Heath, A. C., Meyer, J., Jardine, R., and Martin, N. G. (1991b). The inheritance of alcohol consumption patterns in a general population twin sample: II. Determinants of consumption frequency and quantity consumed. *J. Stud. Alcohol* **52**: 425–433.

Hopper, J. L., White, V. M., Macaskill, G. T., and Hill, D. J. (1992). Alcohol use, smoking habits and the Junior Eysenck Personality Questionnaire in adolescent Australian twins. *Acta Genet. Med. Gemmellol. Twin Res.* **41**:311–324.

Johnston, L. D., O'Malley, P. M., and Bachman, J. G. (1998). *National Survey Results from the Monitering the Future Study, 1975–1997, Vol. 1. Secondary School Students,* U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.

Kendler, K. S., Gardner, C., and Prescott, C. A. (1997). Religion, psychopathology and substance use and abuse: A multi-measure, genetic-epidemiologic study. *Am. J. Psychiatry* **154**: 322–329.

Kendler, K. S., Neale, M. C., Sullivan, P. F., Corey, L. A., Gardner, C. O., and Prescott, C. A. (1999). A population-based twin study in women of smoking initiation and nicotine dependence. *Psychol. Med.* **29**:299–308.

Koopmans, J. R. (1997) *The Genetics of Health-Related Behaviors. A Study of Adolescent Twins and Their Parents,* Thesis, Vrije Universiteit, Amsterdam.

Koopmans,. J. R., and Boomsma, D. I. (1996). Familial resemblances in alcohol use: Genetic or cultural transmission? *J. Stud. Alcohol* **57**:19–28.

Koopmans, J. R., Boomsma, D. I., Heath, A. C., and van Doornen, L. J. P. (1995). A Multivariate genetic analysis of sensation seeking. *Behav. Genet.* **25**:349–356.

Koopmans, J. R., van Doornen, L. J. P., and Boomsma, D. I. (1997). The association between alcohol use and smoking in adolescent and young adult twins: A bivariate genetic analysis. *Alcohol. Clin. Exp. Res.* **21**(3):537–546.

Maes, H. M., Woodard, C. E., Murrelle, L., Meyer, J. M., Silberg, J. L., Hewitt, J. K., Rutter, M., Simonoff, E., Pickles, A., Carbonneau, R., Neale, M. C., and Eaves, L. J. (1999). Tobacco, alcohol and drug use in eight- to sixteen-year-old twins: The Virginia Twin Study of Adolescent Behavioral Development. *J. Stud. Alcohol* **60**:293–305.

Magnus, P., Berg, K., and Nance, W. E. (1983). Predicting zygosity in Norwegian twin pairs born 1915–1960. *Clin. Genet.* **24**: 103–112.

McGue, M., Sharma, A., and Benson, P. (1996). Parent and sibling influences on adolescent alcohol use and misuse: Evidence from a U.S. adoption cohort. *J. Stud. Alcohol* **57**(1):8–18.

Neale, M. C. (1995). *Mx: Statistical Modeling, 3rd ed.,* Department of Psychiatry, Box 710 MCV, Richmond, VA23298.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families,* Kluwer Academic, Dordrecht.

Neale, M. C., Eaves, L. J., Kendler, K. S., and Hewitt, J. K. (1989). Bias in correlations from selected samples of relatives: The effects of soft selection. *Behav. Genet.* **19**:163–169.

Office of Applied Studies (1999). *National Household Survey on Drug Abuse: Main Findings 1997,* Department of Health and Human Services, Substance Abuse and Mental Health Services Administration.

Pianezza, M. L., Sellers, E. M., and Tyndale, R. F. (1998). Nicotine metabolism defect reduces smoking. *Nature* **393**:750.

Pomerleau. (1995). Individual differences in sensitivity to nicotine: Implications for genetic research on nicotine dependence. *Behav. Genet.* **25**(2):161–177.

Rietveld, M. J. H., Koopmans, J. R., Maes, H. H., and Boomsma, D. I. (1996). Genetic and environmental influences on alcohol use and smoking by maternal religious involvement in Dutch adolescent twins. *Behav. Genet.* **26**(6):595 (abstract).

True, W. R., Heath, A. C., Scherrer, J. F., Waterman, B., Goldberg, J., Lin, N., Eisen, S. A., Lyons, M. J., and Tsuang, M. T. (1997). Genetic and environmental contributions to smoking. *Addiction* **92**(10):1277–1287.

Urberg, K. A., Degirmencioglu, S. M., and Pilgrim, C. (1997). Close friend and group influence on adolescent cigarette smoking and alcohol use. *Dev. Psychol.* **33**(5):834–844.

World Health Organization (1997). *Tobacco or Health: A Global Status Report,* WHO, Geneva.

Zuckerman, M. (1994). *Behavioral Expressions and Biosocial Bases of Sensation Seeking,* Cambridge University Press, New York.

Edited by Richard J. Rose